How Does Duration Between Trades of Underlying Securities Affect Option Prices*

Álvaro Cartea†

Birkbeck, University of London, UK

Email: a.cartea@bbk.ac.uk

Thilo Meyer-Brandis

CMA, University of Oslo

Email: meyerbr@math.uio.no

November 26, 2007

*First Version April 2004. Corresponding author a.cartea@bbk.ac.uk. This paper has benefited from comments of seminar participants at The University of Chicago, Federal Reserve Bank Chicago, University of Florence, University of Toronto, University of Oxford, ESSEC, King's College London and Birkbeck-University of London. For comments and suggestions on earlier drafts, we are indebted to Gene Amromin, Luca Benzoni, Raymond Brummelhuis, Marcelo G. Figueroa, Craig Furfine, Helyette Geman, Sam Howison, Aytac Ihlan, Sebastian Jaimungal, Tim Jenkinson, Pete Kyle, David Marshall, Dilip Madan, Colin Mayer, Robert McDonald, Roel Oomen, Andrea Roncoroni, Oren Sussman and Jos van Bommel.

†Á. Cartea is thankful for the hospitality and generosity shown by the Finance Group at the Saïd Business School, Oxford, where part of this research was undertaken.

ABSTRACT

We propose a model for stock price dynamics that explicitly incorporates random waiting times between trades, also known as duration, and show how option prices can be calculated using this model. We use ultra-high-frequency data for blue-chip companies to motivate a particular choice of waiting-time distribution and then calibrate risk-neutral parameters from options data. We also show that the convexity commonly observed in implied volatilities may be explained by the presence of duration between trades. Furthermore, we find that, ceteris paribus, implied volatility decreases in the presence of longer durations, a result consistent with the findings of Engle (2000) and Dufour and Engle (2000) which demonstrates the relationship between levels of activity and volatility for stock prices.

Keywords: Duration between trades, waiting-times, high frequency data, Lévy processes, option pricing, time changes, operational time, irregularly spaced data.

Most financial models assume that securities are continuously traded. However, in equity markets for example, trading happens discretely at random times. In the literature there have been several approaches to directly model the times between trades also known as duration. Early models that capture the impact of duration between trades include Diamond and Verrechia (1987) and Easley and O'Hara (1992). The work of Easley and O'Hara establishes the link between the existence of information, the timing of trades and the dynamics of security prices. One of their main contributions is to show that duration between trades affects the behavior of security prices and consequently that transaction prices are not a Markov process, as is currently assumed in many financial models.

Using ultra-high-frequency equity data, Engle (2000) studies the consequences of stochastic trade arrival times (see also Engle and Russell (1998)). This empirical study finds evidence that both stock returns and variances are found to be negatively influenced by long durations between trades. The study of Dufour and Engle (2000) shows that the stochastic component of duration can explain the relationship between short time durations, i.e. high trading activity, and both larger quote revisions and stronger positive autocorrelations of trades.

Recent work by Aït-Sahalia and Mykland (2003) focuses on the estimation of continuous-time models and its consequences, in particular the fact that high-frequency financial data are discretely sampled in time and that the time separating successive observations is often random. One of the main messages emerging from their findings is that for empirical purposes, researchers using randomly spaced data, "... should pay as much attention, if not more, to sampling randomness as they do to sampling discreteness".

When it comes to derivative pricing, most financial literature on discrete time models assumes that the distribution of the waiting-time $\tau_n = T_n - T_{n-1}$ between the nth and $(n-1)$th trades, occurring at times T_n and T_{n-1} respectively, is either constant (tree models) or exponentially distributed (compound Poisson process models). This prompts two questions. Firstly, to what extent are these assumptions deviating from the 'true' distribution of durations? Secondly, how will this deviation from the 'true' empirical distribution impact derivative prices?

The first question is not a new line of research in the literature, but the second, despite its importance in asset pricing, has not been addressed until now.

When looking at data that involves the random arrival of events, trades in our case, it is customary to look at what is known as the survival function, which represents the probability that the waiting-time between two consecutive trades is greater than t. This function is given by

$$\Upsilon(t) = 1 - \int_0^t \upsilon(u)du, \tag{1}$$

where $\upsilon(t)$ denotes the probability density function (pdf) of the waiting times.

If we assume that the waiting-time between trades possesses an exponential distribution with parameter λ, then $\upsilon(t) = \lambda e^{-\lambda t}$ and $\Upsilon(t) = e^{-\lambda t}$. Employing General Motors (GM) consolidated trades (over the period April-June 2005) in Figure 1, as an example we show a log-log plot of empirical and fitted exponential survival functions.[1] We used 419,264 trades from all exchanges with a resulting average duration between consecutive trades of $\tau_o^e = 5.26$ seconds. The Figure also shows that the fitted exponential survival function with parameter $\lambda = 1/\tau_o^e$, (the dashed line), is a very poor fit when compared to empirical data (circles).[2]

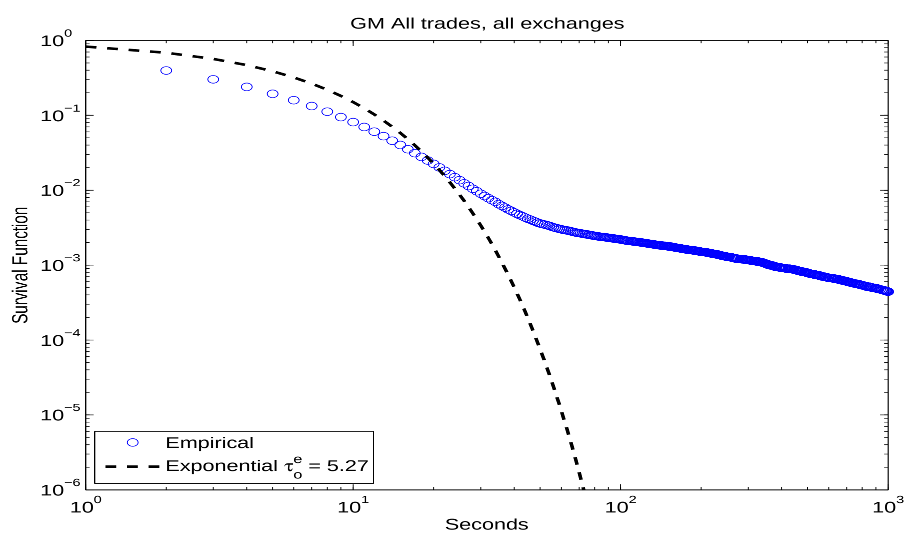

Figure 1. General Motors waiting-times: empirical and exponential.

Intuitively, the rationale for rejecting the exponential survival function as a possible candidate to model durations is its inability to capture the long durations between consecutive trades, see for example Engle (2000), Engle and Russell (1998) and Dufour and Engle (2000). Furthermore, assuming that the duration between consecutive trades is exponentially distributed is equivalent to assuming that the number of trades follows a Poisson counting process. If this were the case, then the mean and variance of the data should be the same, a property know as 'equidispersion'. In fact, what is commonly observed in ultra-high-frequency models is 'overdisperion', i.e. where the variance is greater than the mean of the data, see Cameron and Trivedi (1996). For example, in the case of GM the variance of waiting times for trades is $3.4575 * 10^3$, while the mean is 5.27.

In this article, we concentrate on the question of how derivatives prices are calculated when durations possess a distribution function that better reflects the observed empirical behavior. Our contribution is threefold. Firstly, we propose a general model that explicitly incorporates waiting times as one of the building blocks of stock price dynamics under the physical measure. Secondly, we show how option prices are calculated by choosing a risk-adjusted measure. Thirdly, based on empirical waiting-time data from blue-chip companies, we investigate a particular distribution for duration and we employ it to calibrate risk-neutral parameters to IBM options data.

Under the risk-adjusted measure we propose the use of a survival function that can capture long waits between trades and that nests, as a particular case, the exponential survival function. We then calibrate our model to IBM options data and find that in the vast majority of the cases the risk-neutral parameters of the stock dynamics responsible for modeling the duration between trades, indicate that the risk-neutral distribution of waiting times is not exponential.

As another illustration of our model, we chose to isolate the effect of the waits by calculating option prices with a survival function that allows for the possibility of long waits and the distribution of stock price revisions can be either Gaussian or CGMY (see Carr, Geman, Madan, and Yor (2002)). We observe that for different maturities the inclusion of waiting-times that are not exponentially distributed contribute to the implied volatility observed in financial markets. In particular, when we assume that price revisions are Gaussian, which asymptotically behaves like the classical Black-Scholes framework, the inclusion of non-exponential waiting-times is solely responsible for the emergence of the convexity in the volatility 'smile'. We also observe that, ceteris paribus, implied volatility decreases when waiting times are 'longer', a finding in line with those of Engle (2000) and Dufour and Engle (2000) which links the relationship of levels of activity and volatility for stock prices.

The rest of this article is organized as follows. Section I proposes a general model for stock prices, under the statistical measure, where duration between trades is random. Section II focuses on the pricing of instruments such as European-style options. Section III justifies the selection of particular waiting-time distributions and shows how European-style option prices may be calculated by employing widespread techniques such as those in Carr and Madan (1999). Section IV calibrates risk-neutral parameters for one of our models, using IBM options data. Section V produces numerical examples of how duration affects the shape and level of implied volatility. Section VI concludes.

I. The Model: spot dynamics with duration

In this section, we propose a model which needs to satisfy three requirements. Firstly, every time a trade occurs stock prices must undergo a stochastic price revision. Secondly, the model must be able to explicitly incorporate the (random) duration between trades. Thirdly, the model must be capable of pricing basic financial instruments such as European-style options and one must be able to calibrate its risk-neutral parameters to the market.

Before presenting the model we need two more definitions: a counting process; and the hazard function. We denote the time of the nth trade by T_n and the duration between trades by $T_n - T_{n-1} = \tau_n$ with continuous pdf $\upsilon(t)$. Hence we can write

$$T_n = T_0 + \sum_{i=1}^{n} \tau_i, \qquad T_n - T_{n-1} = \tau_n, \qquad n = 1, 2, 3, \cdots .$$

The counting process, which represents the number of trades over the interval $[0, t]$, is defined by

$$N_t = \max\{n \geq 0 | T_n \leq t\} .$$

Further, the hazard function $u(t)$ is defined as

$$u(t) = -\frac{d}{dt} \ln \Upsilon(t) , \qquad t \in \mathbb{R}^+, \tag{2}$$

where the survival function $\Upsilon(t)$ is that given above in equation (1). Intuitively, the hazard function represents the probability that a trade will happen in the next small time interval divided by the length of that time interval; i.e. the hazard function is the instantaneous intensity of a trade occurrence. Here we assume that $u(t)$ is strictly positive and continuous.

Stock price revisions

To model the stock price revisions, we assume that every time there is a trade, i.e. the counting process N_t increases by one unit, the price revision of the logarithm of the stock price $X(t) = \ln S(t)$ moves by i.i.d. Y. More precisely, we assume that the dynamics of the observed tick-by-tick microstructure of $X(t)$, under the physical measure $\mathbb{P}$, are described by

$$X(t) = X(0) + (r - D)t + \sum_{i=1}^{N_t} Y_i , \tag{3}$$

where the constants r and D denote the risk-free rate and the dividend yield. Note that for technical convenience, we consider a continuously compounded risk-free behavior with rate $(r-D)$ instead of capturing this deterministic trend in the jump price revisions $\sum_{i=1}^{N_t} Y_i$. At jump times (i.e. when there is a trade) there is no price difference between these two alternatives. However, with the continuous rate technicalities are simplified when it comes to derivatives pricing in section II below. We assume that the i.i.d. spacial shocks Y, which are independent of the waiting times, possess an infinitely divisible distribution. Given the above, the log-characteristic function of Y is given by the Lévy-Khintchine representation

$$\ln \mathbb{E}\left[e^{i\xi Y_i}\right] \equiv \Psi(\xi) = ai\xi - \frac{1}{2}\sigma^2\xi^2 + \int_{\mathbb{R}\setminus\{0\}} \left(e^{i\xi l} - 1 - i\xi l 1_{|l|<1}\right) W(dl). \tag{4}$$

Here $a \in \mathbb{R}$, $\sigma \geq 0$, the truncation function $l1_{|l|<1}$ ensures integrability around the origin, and $\Psi(\xi)$ is known as the characteristic exponent of the distribution with triplet (a, σ^2, W). For technical simplicity, we assume that the distribution of the spacial shocks Y is given by a continuous density $g(y) > 0$, $y \in \mathbb{R}$. Note that if we denote by $N(\omega, dt, dz) = N(dt, dz)$ the integer valued jump measure associated with the process $\sum_{i=1}^{N_t} Y_i$, we can rewrite the dynamics (3) as[3]

$$X(t) = X(0) + (r-D)t + \int_0^t \int_{\mathbb{R}_0} zN(dt, dz)\,. \tag{5}$$

In the financial literature, the two most common models of the type described in equation (3) are: discrete time models (tree models) with deterministic, equally spaced, time steps τ_n; and compound Poisson models where the τ_n's are i.i.d. exponentially distributed, random variables. In the latter, $X(t)$ belongs to the class of Lévy processes which have been extensively studied and applied in finance over the recent years.

For example, a conditionally Gaussian model arises when it is assumed that price revisions in (3) arise from a Gaussian distribution, with $Y \sim N(\mu, \sigma^2)$, and that the counting process N_t is a homogeneous Poisson process, which is equivalent to assuming that the waiting-time distribution between trades is exponential. However, as is well known, the Gaussianity of price revisions is not supported by empirical studies, especially over short-time periods. Most efforts to improve these models have focused on the spacial shocks aspect, as opposed to the distribution of the waiting times τ, despite the crucial role that these waiting times play in the distributional properties of stock prices.

A major reason why people only reluctantly depart from exponentially distributed waiting times, is the loss of Markovianity (even if empirical studies confirm the non-Markovianity of prices). Indeed Markovianity is important for many issues, including derivatives pricing, where expectations conditioned on past market evolution have to be computed. With the exception of the exponential waiting-time distribution, the log-stock $X(t)$ is not Markovian for a general waiting-time distribution in model (3). Indeed, let $H(\omega,t) = H(t)$ denote the so-called backward recurrence time (i.e. the time elapsed since the last trade) defined by

$$H(t) = t - T_{N_t}\,, \tag{6}$$

where T_{N_t} represents the last trade time before t. Then it is well known (see e.g. Jacobsen (2006)) that the intensity of the counting process N_t is given by $u(H(t))$ Consequently, the predictable compensator of the jump measure $N(dt,dz)$ is the random measure

$$\nu(\omega,dt,dz) = \nu(dt,dz) := u(H(t))g(z)\,dtdz\,, \tag{7}$$

where $u(t)$ was the hazard function given in (2) and $g(z)$ the probability density of the shocks Y. From this it follows that the process is not Markovian as long as $u(t)$ is not constant. Intuitively, for general hazard functions $u(t)$, it is important to know the time elapsed since the last trade and thus the process is not memoryless. However, if we enlarge the state space with the backward recurrence time $H(t)$, then we have the following result.

Theorem 1 *The two dimensional process $(X(t),H(t))$ is a time-homogeneous Markov process.*

This is an important property which we will use below to price options. For a proof see appendix A.

A special example is the well-known case resulting from the assumption that the waiting times τ are exponentially distributed with parameter λ. For this particular case, the survival function is given by $\Upsilon(t) = e^{-\lambda t}$ and the hazard function becomes $u(t) = \lambda$; note that the hazard function is independent of the backward recurrence time $H(t)$. In this case the compensating measure (7) becomes $\nu^2(\omega, dt, dz) = \lambda g(z)\, dt dz$, which is the compensating measure of the compound Poisson process $X(t)$, and it is not necessary to consider the two-dimensional process $(X(t), H(t))$ because $X(t)$ already is Markovian.

II. Derivatives Pricing

One of the key requirements we have imposed on our model for stock price dynamics is that we can price financial instruments, such as European-style options written on the underlying stock $S(t)$. Therefore, in the first part of this section, we discuss the possible risk-neutral dynamics exhibited by $S(t)$ when we assume that, under the physical measure $\mathbb{P}$, the stock price follows (3). In the second part we then proceed to discuss derivatives pricing and derive an integro-pde characterization for the price process of European-style options in our framework. Further, under the assumption that a trade just has happened, we derive a second price description based on Fourier transform techniques which is much more efficient in practice both to price, and more importantly, to calibrate risk-neutral parameters.

On our stochastic basis $(\Omega, \mathcal{F}, \mathbb{P})$, let $\mathcal{F}_t$ be the filtration generated by the stock price $S(t)$; note that the same filtration is generated by the two-dimensional process $(X(t), H(t))$. Since $S(t)$ is obviously a semimartingale, theory tells us that we must specify an equivalent martingale measure (EMM) $\mathbb{Q}$, under which risk-neutral pricing of financial instruments, written on $S(t)$, can be performed. Given the market incompleteness in our model there is no unique EMM and it is the market that chooses an EMM under which prices are computed.

To specify the family of potential EMMs, we adopt the same approach employed by the vast majority of incomplete market models. We assume that the stock dynamics under the risk-adjusted measure have the same structure as under the physical measure. For example, in the Lévy process literature it is assumed that stock prices will follow a Lévy process under both the physical and risk-neutral measure, but not necessarily the same one (Cont and Tankov (2004)).[4] Therefore, we will assume that the risk-neutral process will possess the same structure under both measures. In particular, the number of trades will be independent from price revisions, but we allow the distribution of the number of trades to be different under the risk-neutral measure. In addition the distribution of price revisions is again infinitely divisible, but not necessarily the same one as under the physical measure.

More precisely, we assume that the market chooses from a class of EMMs whose densities with respect to $\mathbb{P}$ is given by the following stochastic exponentials

$$\frac{d\mathbb{Q}}{d\mathbb{P}} = \exp\left(\int_0^t \int_{\mathbb{R}_0} \ln(\phi(z)\alpha(\omega,t))\, N(dt,dz) - \int_0^t \int_{\mathbb{R}_0} (\phi(z)\alpha(\omega,t)-1)\, \nu(dz,dt)\right), \qquad (8)$$

where the function $\phi(z)$ and the predictable process $\alpha(\omega,t)$ are such that (8) is a well defined $\mathbb{P}$-martingale. Further, we assume that $g_{\mathbb{Q}}(z) = \phi(z)g(z)$ is the density of an infinitely divisible distribution satisfying

$$\int_{\mathbb{R}} (e^z - 1) g_{\mathbb{Q}}(z)\, dz = 0\,, \qquad (9)$$

and that $\alpha(\omega,t)u(H(t))$ takes the form $u_{\mathbb{Q}}(H(t))$ for a strictly positive and continuous hazard function $u_{\mathbb{Q}}(t)$. Using Girsanov's theorem for random measures (see Jacod and Shiryaev (2002)), the jump measure $N(\omega,dz,dt)$ has the $\mathbb{Q}$-predictable compensator

$$\nu_{\mathbb{Q}}(\omega,dt,dz) = u_{\mathbb{Q}}(H(t))g_{\mathbb{Q}}(z)\, dtdz\,, \qquad (10)$$

which has the same structure as the predictable compensator (7) under the $\mathbb{P}$ measure. It is straightforward to see from the structure of the $\mathbb{Q}$-compensator (10) that the log-stock price

$$\begin{aligned} X(t) &= X(0) + (r-D)t + \sum_{i=1}^{N_t} Y_i \\ &= X(0) + (r-D)t + \int_0^t \int_{\mathbb{R}_0} z N(dt, dz) \end{aligned}$$

has the same renewal process structure under $\mathbb{Q}$, as it has under $\mathbb{P}$. The alteration is only a different, but equivalent infinitely divisible distribution for the spacial shocks Y given through the density $g_Q(z)$, which is such that $\mathbb{E}^{\mathbb{Q}}[e^Y - 1] = 0$, as well as a different hazard function $u_Q(t)$ characterizing the waiting times. Now, the discounted stock price $e^{-(r-D)t}S(t)$ is given by

$$e^{-(r-D)t}S(t) = S(0)\exp\left(\int_0^t \int_{\mathbb{R}_0} z N(dt, dz)\right).$$

Because of condition (9) we can rewrite $e^{-(r-D)t}S(t)$ as

$$e^{-(r-D)t}S(t) = S(0)\exp\left(\int_0^t \int_{\mathbb{R}_0} z N(dt, dz) - \int_0^t \int_{\mathbb{R}_0} (e^z - 1)\,\nu_Q^2(dz, dt)\right), \tag{11}$$

which is an exponential martingale under $\mathbb{Q}$. Consequently, under the above conditions, (8) determines indeed a class of EMM.

Having specified a pricing measure $\mathbb{Q}$ from the above defined class, we now consider pricing of instruments written on $S(t) = \exp(X(t))$. Let F be a pay-off function of a European option with maturity T written on $S(t)$. Then the price process of this option is given as

$$V(t) = e^{-r(T-t)}\mathbb{E}^{\mathbb{Q}}[F(S(T))|\mathcal{F}_t] \qquad 0 \le t \le T\,.$$

Note that considering a European option written on $S(t)$ is equivalent to considering a European option written on $X(t)$ with pay-off function $G = F(\exp(\cdot))$. Thus, the value process of the option above can be rewritten as

$$V(t) = e^{-r(T-t)}\mathbb{E}^{\mathbb{Q}}[G(X(T))|\mathcal{F}_t] .$$

Now, because of the time-homogeneous Markov structure of $(X(t),H(t))$, we can write

$$V(t) = e^{-r(T-t)}\mathbb{E}^{\mathbb{Q}}[G(X(T))|X(t),H(t)] = e^{-r(T-t)}\mathbb{E}^{x}_{\mathbb{Q}}[G(X^h(T-t))] \,|_{x=X(t),h=H(t)} . \quad (12)$$

Here, $X^h(t)$ is the h-delayed renewal process starting in x, induced by $X(t)$, i.e. the first waiting-time in (3) has the distribution of $(\tau_1 - h)$, given $\tau_1 > h$. Furthermore, from (A2) and (A3) it follows that the generator of the Markov process $(X(t),H(t))$ is given through the integro-differential operator $\mathcal{O}$, defined as follows:

$$\mathcal{O} f(x,h) = (r-D)\frac{\partial}{\partial x} f(x,h) + \frac{\partial}{\partial h} f(x,h) + \int_{\mathbb{R}_0} \{f(x+z,0) - f(x,h)\} u_Q(h) g_Q(z)\, dz, \quad (13)$$

for $f \in C_0^{1,1}(\mathbb{R}^2)$. Here, $C_0^{1,1}(\mathbb{R}^2)$ is the space of continuous functions, with compact support and continuous derivatives in x and h. Then, with the usual Feynman-Kac considerations, we obtain the following description of the price process $V(t)$.

Theorem 2 *Let $F(\cdot)$ be the pay-off function of a European option with maturity T written on the stock $S(t)$. Let the function $G(\cdot) := F(\exp(\cdot))$ be the composition of F and* exp*, and assume that there exists a bounded solution $v(t,x,h) \in C^{1,1,1}([0,T],\mathbb{R},\mathbb{R}_+)$ of the integro-pde*

$$\begin{cases} 0 = \frac{\partial}{\partial t} v(t,x,h) + \mathcal{O} v(t,x,h) \\ \\ v(T,x,h) = G(x), \ \ (t,x,h) \in [0,T] \times \mathbb{R} \times \mathbb{R}_+. \end{cases} \quad (14)$$

Then, the price at time t of the European option with pay-off $F(\cdot)$, and maturity T, is given as

$$V(t) = e^{-r(T-t)} v(t, X(t), H(t)).$$

Note that in the special case of an exponential waiting time distribution with parameter λ, the generator (13) becomes

$$\mathcal{O} f(x,h) = (r-D)\frac{\partial}{\partial x} f(x,h) + \frac{\partial}{\partial h} f(x,h) + \int_{\mathbb{R}_0} \{f(x+z,0) - f(x,h)\} \lambda g_Q(z)\, dz.$$

Thus, if a function $v'(t,x) \in C^{1,1}([0,T],\mathbb{R})$ solves

$$\begin{cases} 0 = \frac{\partial}{\partial t} v'(t,x) + \mathcal{O}' v'(t,x) \\ \\ v'(T,x) = G(x), \;\; (t,x) \in [0,T] \times \mathbb{R}, \end{cases} \tag{15}$$

where the generator $\mathcal{O}'$ is defined as

$$\mathcal{O}' f(x) = (r-D)\frac{\partial}{\partial x} f(x) + \int_{\mathbb{R}_0} \{f(x+z) - f(x)\} \lambda g_Q(z)\, dz,$$

$f \in C_0^1(\mathbb{R})$, then $v(t,x,h) := v'(t,x)$ solves (14). Consequently, for exponentially distributed waiting times, we obtain the usual pricing integro-pde (15) for compound Poisson processes which is independent of h.

The integro-pde representation of the option price (14) provides a method for computing option prices in our model. However, an alternative way to calculate prices of European-style instruments is to use transform methods (Carr and Madan (1999), Carr and Wu (2003)). These methods are very efficient and powerful to calibrate risk-neutral parameters from market data. Here we present the general result which we employ below in subsection A.1, when we choose a particular survival function, to calibrate parameters to IBM options data in Section IV.

Proposition 1 *Let $F(\cdot)$ be the pay-off function of a European option with maturity T written on the stock $S(t)$, and let $G(\cdot)$ be as in Theorem 2. Assume that $\hat{q}(\xi,t,T)$, defined by*

$$\hat{q}(\xi,t,T) := \mathbb{E}^Q\left[e^{i\xi\sum_{i=N_t+1}^{N_T} Y_i} \mid \mathcal{F}_t\right], \tag{16}$$

is analytic in ξ in a strip that intersects the strip where the (complex) Fourier transform of G exists. Let $\hat{\xi} \in \mathbb{R}$ be such that the line $[-\infty + i\hat{\xi}, \infty + i\hat{\xi}]$ is part of this intersection. Then the value at time t of the option is given by

$$V(t) = \frac{e^{-r(T-t)}}{2\pi}\int_{-\infty+i\hat{\xi}}^{\infty+i\hat{\xi}} e^{-i\xi \ln S(t)} e^{-i\xi(r-D)(T-t)} \hat{q}(-\xi,t,T)\hat{G}(\xi)d\xi. \tag{17}$$

where the notation $\hat{G}(\xi) = \mathcal{F}[G(x)] = \int_{-\infty}^{\infty} e^{ix\xi}G(x)dx$ denotes the Fourier transform of $G(\cdot)$.

For a proof see appendix A.

We note that, depending on the assumptions regarding the waiting-time distribution $v(t)$, and/or the counting process N_t, expression (16) can be calculated analytically and the evaluation of European-style option prices becomes a straightforward task.

III. Empirical survival function

In this section we look at empirical waiting-times of 23 blue-chip companies during the period April-June 2005. Our sample of stocks includes those from Dufour and Engle (2000) that were still being traded in 2005. All data were obtained from the TAQ database made available via WRDS.

Before proposing a model that captures the main properties of the empirical survival functions we address the question of how to treat the relatively frequent occurrences of consecutive trades when the duration between them is reported in the system with zero. From a practical point of view, time-stamps for every trade are rounded to the nearest second. A direct consequence of this is that trades that occur within the same second are recorded as if they had taken place simultaneously. On the other hand, there are cases when one trade is broken into various batches and these too are recorded as simultaneous trades. A common approach adopted in the literature has been to delete these trades. For instance, in our data set of IBM trades there are 178,512 durations of zero seconds. Deleting these observations would amount to discarding more than 28% of the 631,586 waits between trades.

Ideally, if one could discern which zero-duration trades are part of a large trade broken into batches, then these could be deleted and the remaining zero-duration trades could be kept by assigning them a waiting-time strictly greater than zero. From a mathematical standpoint, if we view the question of modeling durations as modeling the number of trades occurring on a given interval, we know that counting processes such as Poisson will assign zero probability to events where two or more trades take place at the same time. Therefore the need to assign waiting times that occurred within a second, but recorded as simultaneous trades, a duration strictly greater than zero. Instead of discarding all zero-duration observations the alternative we propose is to remove only those data points where there was a zero waiting-time but there was no change in the price of the trade. For example, of the 178,512 instance of zero-duration in the restricted IBM data, 103,391 could be eliminated because they were accompanied by no change in price. The remaining 75,121 data points where price changes were different from zero were retained and were assigned a duration strictly greater than zero.[5] In Table I, we show, for each stock, the number of data points omitted due to zero waiting times and no price changes (column "Out") and those included through assignment of a non-zero waiting-time (column "In").

Co	Out	In	All Trades	τ_o^e
GE	227,431	84,404	620,370	3.96
IBM	103,391	75,121	528,195	4.27
GM	115,967	61,966	419,264	5.27
MO	63,480	34,527	364,331	5.98
PG	60,038	29,458	365,800	5.54
AMD	89,449	30,209	333,248	6.59
SLB	48,283	30,200	356,341	5.41
KO	53,113	23,066	342,880	5.61
BA	52,328	26,201	323,436	6.12
AA	47,733	19,267	298,566	6.43
FNM	39,579	22,055	296,854	6.13
FDX	3,0545	21,407	260,044	7.31
CL	23,235	9,948	201,127	8.93
FPL	16,015	10,344	188,586	9.27
CAL	22,243	5,895	164,403	10.94
CAG	14,707	7,674	167,293	10.71
T	13,892	5,249	156,005	11.58
PCO	10,159	7,640	155,465	11.56
VC	18,366	6,756	130,115	14.45
HNZ	10,552	3,242	132,931	13.19
NI	8,294	3,144	105,780	16.42
POM	2,407	2,132	69,986	24.51
GTI	3745	979	62,016	27.51

Table I

Empirical waiting-time data. The second column, under the heading "Out", indicates the number of data points, for each stock, that were discarded because a zero wait was also accompanied by a zero price change. The third column, under the heading "In", shows the number of data points which were kept because although there was a zero wait, price changes were not zero. The fourth column indicates therefore the number of data points used as duration between trades. Finally the fifth column is the average waiting time (in seconds) for the data set.

A. Shifted-Mittag-Leffler survival function

The most conspicuous message from Figure 1 is the presence of relatively 'long' durations. These long durations are impossible to capture with an exponential waiting-time distribution, and, as we shall see below, the presence of these long waits between trades is not unique to GM. The appendix shows 22 other companies that exhibit broadly the same shaped survival function as GM. Hence, we will justify a choice of waiting-time distribution by specifying a model that can capture the right tail of the survival function, i.e. long waits.

The first step is to observe that the shape of the right tail of the survival function, in log-log space, in Figure 1 closely resembles that of a straight line with a negative slope. It is straightforward to see that this linear behavior in a log-log plot is equivalent to observing the behavior of data that is changing with a power law. In other words the (ln-)tail of the survival function shows the behavior

$$\ln \Upsilon(t) \sim -\beta \ln t + \ln a + \cdots, \qquad \text{as} \quad t \to \infty, \tag{18}$$

where $\beta > 0$ and a are constants.[6] Since from (1) we obtain the pdf of the waiting times by differentiating the survival function

$$\upsilon(t) = -\frac{d}{dt}\Upsilon(t),$$

we can use (18) to find the tail behavior of the pdf of the waiting-time distribution:

$$\ln \upsilon(t) \sim -(\beta+1)\ln t + \ln(a\beta) + \cdots, \qquad \text{as} \quad t \to \infty. \tag{19}$$

Now that we are able to capture the crucial behavior of long waits via (19), or equivalently via (18), we take the second step and justify the choice of a waiting-time distribution. We recall that we want to be able to use our model for stock dynamics in order to price European-style options. In addition, we would like to specify a waiting-time distribution so that expression (16) in Proposition 1 can be performed analytically.

Instead of working with the tail expression of $v(t)$ given by (19), we look at its Laplace transform. Hence, we can write the tail of the waiting-time distribution in Laplace space as[7]

$$\tilde{\upsilon}(s) \sim 1 - (\tau_o s)^{\beta} + o(s^{\beta}), \qquad \text{for } 0 < \beta \leq 1, \tag{20}$$

where $\tau_o > 0$ is a constant.

However, we are still left with the question of finding a suitable waiting-time distribution since we have only specified the functional form of the tail to capture the long waits. Note that there are many waiting time distributions that could exhibit a slow decay of the right tail, as shown in (20). However not all of them will deliver mathematically tractable expressions capable of being employed by standard pricing tools, and more importantly, will not facilitate the calibration of risk-neutral parameters to observed vanilla option prices (see for example Carr and Madan (1999)). Hence, below we specify $v(t)$ for all $t \geq 0$ by choosing a distribution function that allows us to calculate the characteristic function (16).

We proceed by noting that one possible choice of $\tilde{\upsilon}(s)$, consistent with (20), is given by

$$\tilde{\upsilon}(s) = \frac{1}{1 + (\tau_o s)^{\beta}}, \qquad \text{for } 0 < \beta \leq 1. \tag{21}$$

Moreover, the Laplace transform of the survival function is given by

$$\tilde{\Upsilon}_{ML}(s) = \frac{1 - \tilde{\upsilon}(s)}{s} = \tau_o \frac{(\tau_o s)^{\beta - 1}}{1 + (\tau_o s)^{\beta}}, \qquad \text{for } 0 < \beta \leq 1, \tag{22}$$

and by taking the inverse Laplace transform of (22), see equation (A7) in the appendix, the survival function becomes

$$\Upsilon_{ML}(t) = \sum_{j=0}^{\infty} (-1)^j \frac{(t/\tau_o)^{\beta j}}{\Gamma(\beta j + 1)}, \qquad \text{for } 0 < \beta \leq 1, \tag{23}$$

which is known in the literature as the Mittag-Leffler (ML), or as a generalized, exponential function. Furthermore, we make the important observation that when $\beta = 1$ the waiting-time distribution becomes the exponential with expected value $\mathbb{E}[\tau] = \tau_o$. Hence, we can view the ML survival function as a generalization of the exponential survival function that accommodates long waits between trades when $\beta < 1$; something an exponential waiting-time distribution is unable to capture.

We employ a slight modification of (23), by including a shift parameter τ_s in the time-domain of the survival function. The intuition behind this trivial modification is to recognize that the time-stamps in our data are rounded to the nearest second. Consequently the data set are left-truncated, which therefore makes it reasonable to include a shift in the domain of the survival function to improve the statistical fitting of the ML survival model. Figure 2 shows empirical and fitted survival functions. We show (shifted) ML and exponential functions. As expected, the exponential function is not capable of capturing the long waits. Moreover, Table II shows the results of fitting the shifted ML parameters to all the stocks studied here and the appendix depicts the fitted distributions.

Another route to study empirical waiting times has been to restrict the data set to trading hours between 9.30am and 4.00pm and focus only on trades via NYSE. For example, in this restricted case, the IBM data set would consist of 331,057 trades as opposed to the 528,195 when all exchanges are taken into account and trading before 9.30am and after 4.00pm is also considered. Moreover, previous studies focusing on this restricted data set have found that the Weibull distribution is a good model, however it is not capable of capturing long waits. Moreover, we point out that our main objective is to explicitly model durations and to study their impact on option prices. Therefore, the choice of risk-neutral survival function is what matters when measuring the impact durations have on derivatives pricing.

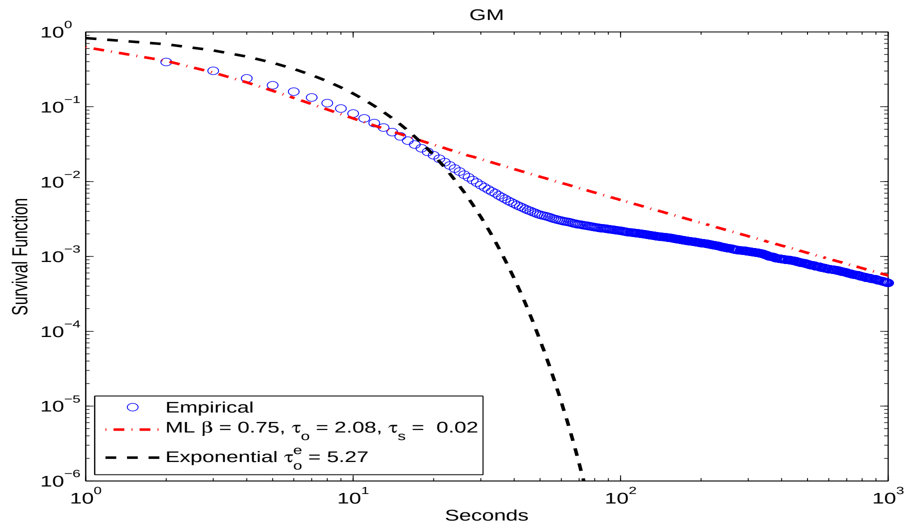

Figure 2. Fitted survival functions for GM

A.1. European-style options with ML survival function

If we assume that, under the risk-neutral measure, the survival function has the form (21) then the problem of pricing European-style options (see Proposition 1) reduces to deriving (16). Furthermore, in this particular case, calculations get simplified if we assume that a trade just happened, i.e $H(0) = 0$, and for simplicity we also assume that $\tau_s = 0$. Given the high frequency of trade arrivals, assuming $H(0) = 0$ is reasonable. The following Theorem shows how European-style options are priced when the survival function of the waiting times is ML.

Theorem 3 *Assume that the prerequisites from Proposition 1 hold. Additionally, assume that the survival function is ML, with* $\tau_s = 0$*, and that* $H(0) = 0$*. Then the value of the European-style option is given by*

$$V(0) = \frac{e^{-rT}}{2\pi} \int_{-\infty+i\hat{\xi}}^{\infty+i\hat{\xi}} e^{-i\xi \ln S(0)} e^{-i\xi(r-D)T} E_{\beta,1}\left[-\left(1-e^{\Psi(\xi)}\right)(T/\tau_o)^{\beta}\right] \hat{G}(\xi) d\xi. \qquad (24)$$

For a proof see appendix A.

Co	τ_o	95% CI	τ_s	95% CI	β	95% CI
GE	1.5060	(1.4717, 1.5403)	0.1103	(0.0998, 0.1208)	0.8423	(0.8293, 0.8554)
IBM	1.7391	(1.6934, 1.7849)	0.0805	(0.0662, 0.0949)	0.8185	(0.8032, 0.8338)
GM	2.0860	(2.0178, 2.1543)	0.0257	(0.0043, 0.0470)	0.7584	(0.7385, 0.7784)
MO	2.6705	(2.6056, 2.7354)	0.0430	(0.0204, 0.0657)	0.8143	(0.8003, 0.8282)
PG	2.8441	(2.7791, 2.9092)	0.0523	(0.0280, 0.0765)	0.8579	(0.8453, 0.8705)
AMD	3.2402	(3.1589, 3.3214)	0.0519	(0.0227, 0.0812)	0.8135	(0.7990, 0.8281)
SLB	2.5343	(2.4663, 2.6022)	0.0833	(0.0608, 0.1059)	0.8000	(0.7841, 0.8159)
KO	2.8949	(2.8278, 2.9621)	0.0747	(0.0505, 0.0990)	0.8398	(0.8266, 0.8530)
BA	2.6259	(2.5444, 2.7074)	0.0773	(0.0516, 0.1030)	0.7556	(0.7363, 0.7750)
AA	3.2311	(3.2065, 3.2556)	0.2206	(0.2179, 0.2233)	0.6452	(0.6408, 0.6497)
FNM	2.8925	(2.7935, 2.9915)	0.0647	(0.0325, 0.0969)	0.7583	(0.7371, 0.7795)
FDX	2.9691	(2.8446, 3.0937)	0.0431	(0.0056, 0.0806)	0.6847	(0.6565, 0.7128)
CL	4.5001	(4.3526, 4.6477)	0.2319	(0.2319, 0.2319)	0.7585	(0.7351, 0.7819)
FPL	4.6416	(4.4736, 4.8096)	0.2349	(0.2349, 0.2349)	0.7351	(0.7086, 0.7616)
CAL	5.2955	(5.1344, 5.4566)	0.2268	(0.2268, 0.2268)	0.7389	(0.7167, 0.7611)
CAG	5.5407	(5.3650, 5.7165)	0.2340	(0.2340, 0.2340)	0.7610	(0.7382, 0.7837)
T	6.1676	(6.0003, 6.3349)	0.2368	(0.2368, 0.2368)	0.7786	(0.7595, 0.7978)
PCO	4.5137	(4.3330 4.6944)	0.2258	(0.2258, 0.2258)	0.6039	(0.5707, 0.6372)
VC	5.8712	(5.6332, 6.1093)	0.2076	(0.2076, 0.2076)	0.6260	(0.5929, 0.6591)
HNZ	7.2854	(7.0743, 7.4964)	0.2345	(0.2345, 0.2345)	0.7791	(0.7585, 0.7997)
NI	9.0244	(8.7679, 9.2809)	0.2409	(0.2409, 0.2409)	0.7573	(0.7366, 0.7780)
POM	14.2969	(13.8032, 14.7907)	0.2439	(0.2439, 0.2439)	0.7518	(0.7262, 0.7775)
GTI	14.7941	(14.3078, 15.2803)	0.2403	(0.2403, 0.2403)	0.7200	(0.6950, 0.7451)

Table II

Shifted ML parameter estimates for τ_o, τ_s (in seconds) and β using ultra-high-frequency data for the trading period April 1st through June 30th 2005.

Regarding the choice of $\hat{\xi}$ in the integration limits in Theorem 3, we require $E_{\beta,1}\left[-\left(1-e^{\Psi(\xi)}\right)(T/\tau_o)^{\beta}\right]$ to be analytic in a strip that intersects the strip where the (complex) Fourier transform of the $G(\cdot)$ exists. The ML function (A6) is an entire function; therefore it is analytic where $e^{\Psi(-\xi)}$ is analytic. Thus, the restrictions on $\hat{\xi}$ are the same as those required in the particular case when $\beta = 1$, i.e. when pricing with Lévy processes.[8] For example, if we let $\beta = 1$, we can verify that the price of a European call option with strike K and maturity T, using (24), is given by

$$V(0;K,T) = -\frac{e^{-rT}K}{2\pi}\int_{-\infty+i\hat{\xi}}^{\infty+i\hat{\xi}} e^{-i\xi\ln S(0)+T\left[-i\xi(r-D)+(\Psi(-\xi)-1)\tau_o^{-1}\right]}\frac{K^{i\xi}}{\xi^2-i\xi}d\xi,$$

for $\hat{\xi} > 1$.[9]

IV. Estimation of risk-neutral parameters

In this section we present results obtained from calibrating risk-neutral parameters to IBM option prices. We obtained data for traded American options written on IBM. This data set include the spot price, strike, maturity, implied volatility, dividend yield and interest rate. We used the parameters from the American options to devise a new data set of European options. We then used the algorithm employed in Carr and Wu (2003) to estimate the risk-neutral parameters of our model by considering two cases. In the first we assume that price revisions possess a Gaussian distribution and that the waiting-time survival function is the ML function. In the second case we still assume that the waiting-time survival function is the ML function but now assume that price revisions possess an FMLS distribution (Carr and Wu (2003)).

The tables in Appendix C show the results for every trading day from April 1 through May 6 2005. In any given day we have IBM options for different strikes and for different maturities. We show the results of the calibration for the lot of IBM options with shortest maturity (including all strikes), then we add to these results the next lot, which includes those options with second shortest maturity, and so on.[10] For example, the first row in Table C shows risk-neutral parameters obtained from 6 options that expired in 10 working days (i.e. the first lot). For this lot, the resulting volatility of Gaussian price revisions and the beta of the model are $\sigma = 0.00058$ and $\beta = 0.717300$ respectively, and for FMLS price revisions $\alpha = 1.99$, $\sigma = 0.000318$ and $\beta = 0.72004$.[11] In the second row, we show the results of the calibration procedure when we take into account the options that expire between 10 and 35 working days.

One of the messages implied by the results is that the effect of long durations (captured by the parameter β) on option prices prevails across all maturities. It is interesting to note that this is true for both the Gaussian and FMLS cases and although the βs are not the same for both models, they do not appear to be too dissimilar for each particular day and surface we calibrate to. We interpret this as a good sign since, especially in the Gaussian example we study, the parameter β could be accommodating for kurtosis of the risk-neutral distribution which, is produced by the spatial shocks, and is 'picked up' by the parameter β. In the next section we see how the presence of long durations ($\beta < 1$) increases the kurtosis of the risk-neutral distribution of spot prices.

V. Numerical examples: the impact of waiting times on option prices

In the previous section, we looked at the calibration of risk-neutral parameters for models that explicitly include waiting times between trades. Here, to gain more insight into the consequences of including durations, we present two examples of how waiting times affect option prices. These are calculated by choosing plausible risk-neutral parameters, so that we can focus on the effects of assuming the ML survival function. The first example assumes that the spatial shocks are Gaussian and the second example assumes that spatial shocks possess a CGMY distribution (see Carr, Geman, Madan, and Yor (2002)). In all examples we assumed that $\tau_o = 1/1,200,000$, (i.e. that there are, on average, 100,000 trades per month) and that $\tau_s = 0$.

A. Gaussian price revisions and ML waiting-times

Figure 3 shows implied volatility (IV) when it is assumed that spatial shocks are Gaussian with mean zero and volatility $\sigma = 0.3\sqrt{\tau_o}$. With this choice of volatility, and letting $\beta = 1$, the model is asymptotically equivalent to assuming a Black-Scholes model with volatility $\sigma_{BS} = 0.30$. The Figure shows IV for different waiting times by choosing $\beta = \{0.98, 0.96, 0.94, 0.92\}$ whilst all other parameters remain unchanged. It is possible to see that the steeper IV becomes for out-of-the-money and in-the-money values the further away the parameter β is from the exponential case $\beta = 1$. This is interesting since it shows that the inclusion of waiting times that are not exponential, gives rise to the commonly observed convexity of the IV in the Black-Scholes framework despite the fact that spatial shocks are Gaussian.[12] Note that the waiting time affects the convexity of the IV in a symmetric way and does not reproduce smirks or skewed IVs. In our framework, market participants include a premium, over and above the classical Black-Scholes price for out-of-the-money values, to price in the duration times between trades.

Another important feature of Figure 3 is the fact that the IV range decreases as β decreases. For example, when $\beta = 0.98$ and expiry is $T = 20$ days, IV is roughly within $[0.265, 0.27]$ whereas when $\beta = 0.94$ and $T = 20$, IV is in the range $[0.21, 0.22]$. This result is not surprising, and is in line with the findings of Engle (2000) and Dufour and Engle (2000). Indeed in our model, the market will exhibit less activity (understood here as number of trades over a time period) and lower IV the lower β is. This is also clear in Figure 4 where, still with Gaussian spacial shocks, we fix expiry dates and vary $\beta = \{1, 0.98, 0.96, 0.94, 0.92\}$ where the exponential case is included.

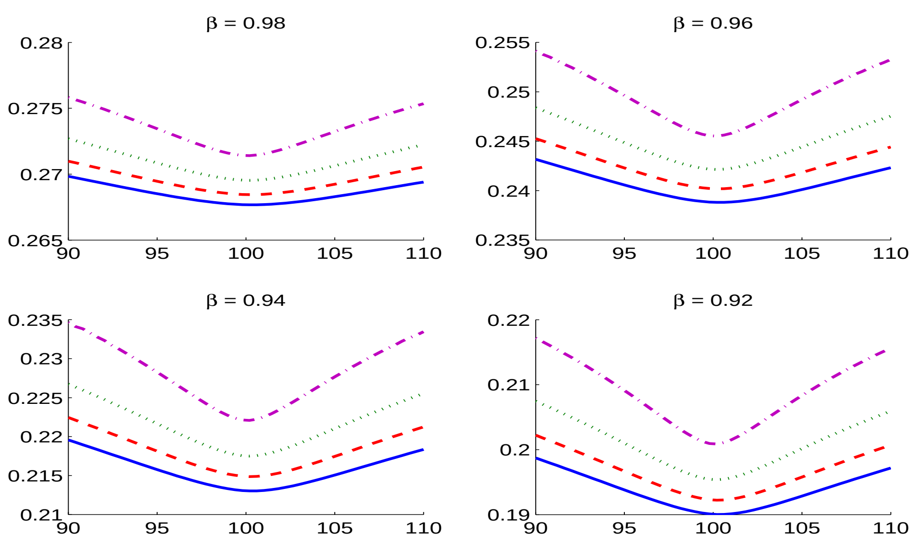

Figure 3. Implied Volatility across strike for conditionally Gaussian model with waiting times for $\beta = \{0.98, 0.96, 0.94, 0.92\}$. The volatility of the zero-mean Gaussian price revisions are $\sigma = 0.3\sqrt{\tau_o}$, and the parameters for option pricing are $r = 5\%$, $D = 0$ and $S_0 = 100$. The dash-dotted line corresponds to $T = 5$ days, the dotted line $T = 10$ days, the dashed line $T = 15$ days, and the solid line $T = 20$ days.

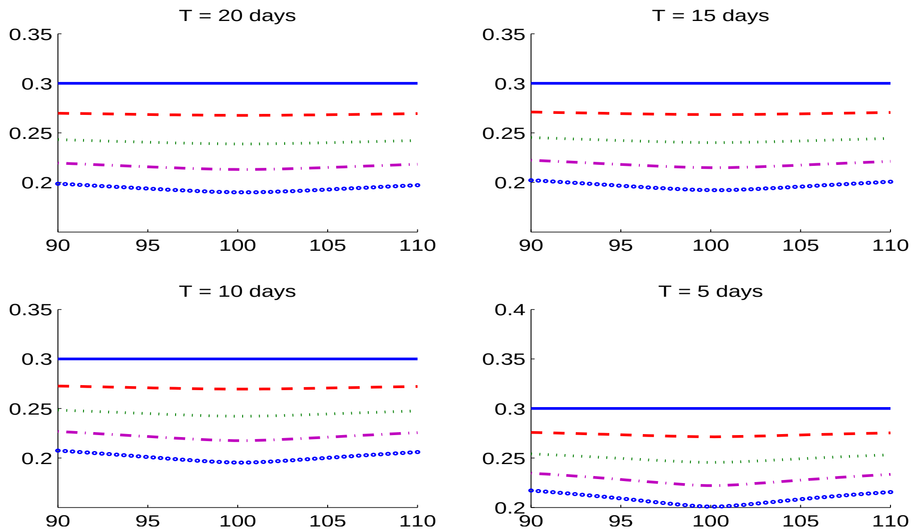

Figure 4. Implied Volatility across strike for conditionally Gaussian with waiting times for different days to maturity $T = \{20, 15, 10, 5\}$ and varying $\beta = \{1, 0.98, 0.96, 0.94, 0.92\}$. The volatility of the zero-mean Gaussian price revisions are $\sigma = 0.3\sqrt{\tau_o}$, and the parameters for option pricing are $r = 5\%$, $D = 0$ and $S_0 = 100$, and the parameters for option pricing are $r = 5\%$, $D = 0$ and $S_0 = 100$. Each panel shows how implied volatility varies when expiry remains fixed and β varies. The solid line represents $\beta = 1$, the dashed line corresponds to $\beta = 0.98$, the dotted line corresponds to $\beta = 0.96$, the dash-dotted line corresponds to $\beta = 0.94$, and circles corresponds to $\beta = 0.92$.

B. CGMY price revisions and ML waiting-times

In this subsection we produce the same results as above, but we allow the distribution of price revisions to exhibit fatter tails than the Gaussian distribution by choosing price revisions with a a CGMY distribution, see Carr, Geman, Madan, and Yor (2002). In our examples below, we assumed that $C = 1.8750 \times 10^{-7}$, $Y = 1.5$, $G = 10$, $M = 20$, this implies that the distribution of the spatial shocks has negative asymmetry because $G < M$, and both the left and right tails of the distribution of spatial shocks are heavier than those of a Normal distribution. We again see the same qualitative results as those from the Gaussian case above.

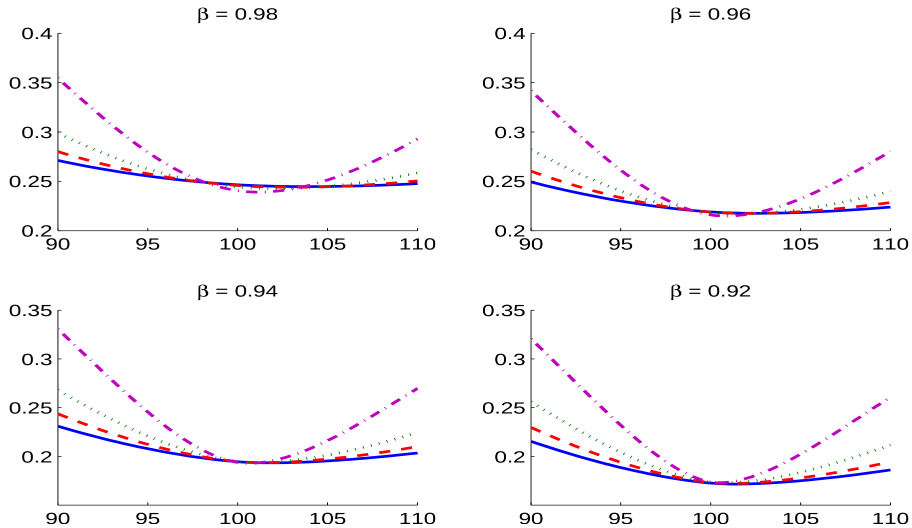

Figure 5. Implied Volatility across strike for CGMY with waiting times for $\beta = \{0.98, 0.96, 0.94, 0.92\}$. The dash-dotted line corresponds to $T = 5$ days, the dotted line $T = 10$ days, the dashed line $T = 15$ days, and the solid line $T = 20$ days.

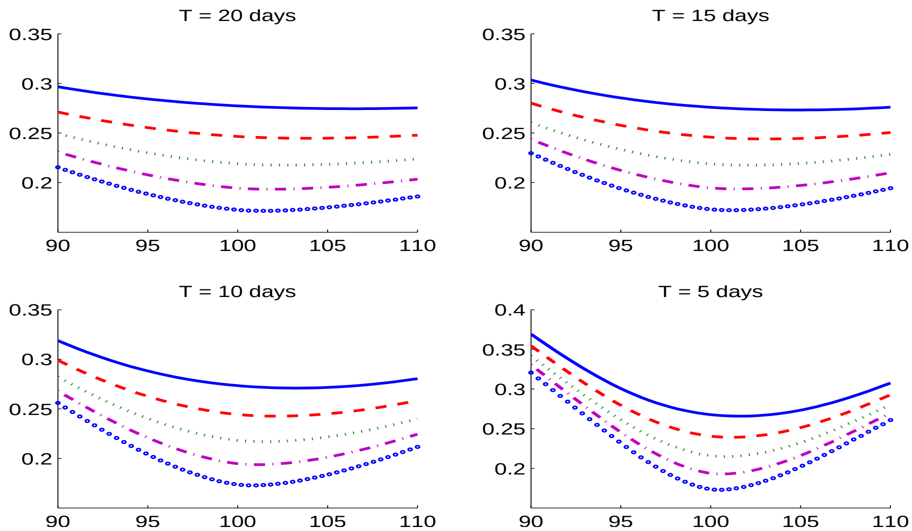

Figure 6. Implied Volatility across strike for CGMY with waiting times for different days to maturity $T = \{20, 15, 10, 5\}$ and varying $\beta = \{1, 0.98, 0.96, 0.94, 0.92\}$. Each panel shows how implied volatility varies when expiry remains fixed and β varies. The solid line represents $\beta = 1$, the dashed line corresponds to $\beta = 0.98$, the dotted line corresponds to $\beta = 0.96$, the dash-dotted line corresponds to $\beta = 0.94$, and circles corresponds to $\beta = 0.92$.

VI. Conclusions

Until now, the financial literature has only considered the question of how waiting-times or duration between trades affect the dynamics of stock prices. The question of how this random duration affects derivative prices, has not previously been addressed. In this article we propose a model that explicitly incorporates these waiting-times. Besides capturing duration between trades, our model also captures key behavioral characteristics recorded in the empirical literature such as the non-Markovianity of stock prices, Easley and O'Hara (1992).

In our model we make the working assumption that waiting-times and spatial shocks are independent. Although this assumption is not endorsed by empirical data, it allows us great flexibility in the modeling of spatial shocks; for example it allows us to assume that price revisions have an infinitely divisible distribution. For this general case, we are able to price European-style options by solving an integro-pde where the standard Lévy-based models (assuming exponentially distributed duration) are a particular case.

We propose the use of the ML survival function as a candidate to model waiting times. One of the main advantages is that with the ML it is straightforward to use the usual transform methods employed in the Lévy process literature relating to finance to price options. As an example, we calibrated risk-neutral parameters, using IBM options data, to a model with ML waits and Gaussian price revisions and to a model with ML waits and FMLS price revisions. In both cases the effects of durations were captured by risk-neutral βs, which were in the vast majority of cases less than one.

As another illustration of our model, we chose to isolate the effect of the waits by calculating options prices with ML waits and Gaussian revision and with ML waits and CGMY price revisions. We saw that for different maturities the inclusion of waiting-times that are not exponentially distributed contribute to the IV observed in financial markets. In particular, when we assume that price revisions are Gaussian, as described by the classical BS framework, the inclusion of waiting-times ($\beta < 1$) is solely responsible for the emergence of the convexity in the volatility 'smile'. Moreover, we see that the level of activity (as the number of trades, over a given time period) is higher the larger β is. We also observe that, ceteris paribus, IV decreases in β a finding in line with those of Engle (2000) and Dufour and Engle (2000) which links the relationship of levels of activity and volatility for stock prices.

References

Aït-Sahalia, Yacine, and Per A. Mykland, 2003, The Effects of random and discrete sampling when estimating continuous-time diffusions, *Econometrica* 71, 483–549.

Cameron, A. Colin, and Pravin K. Trivedi, 1996, Count data models for financial data, in G.S. Maddala, and Calyampudi Rao, eds.: *Handbook of Statistics* (Elsevier, North-Holland, Amsterdam,).

Carr, Peter, Hellyete Geman, Dilip Madan, and Marc Yor, 2002, The fine Structure of asset returns: an empirical investigation, *Journal of Business* 75, 305–332.

Carr, Peter, and Dilip Madan, 1999, Option Valuation Using the Fast Fourier Transform, *Journal of Computational Finance* 2, 61–73.

Carr, Peter, and Liuren Wu, 2003, The Finite Moment Logstable Process and Option Pricing, *The Journal of Finance* LVIII, 753–777.

Carr, Peter, and Liuren Wu, 2004, Time-Changes Lévy Processes and Option Pricing, *Journal of Financial Economics* 71, 113–141.

Cont, Rama, and Peter Tankov, 2004, *Financial Modelling With Jump Processes*. (Chapman and Hall London) 1st edn.

Diamond, Douglas W., and Robert E. Verrechia, 1987, Constraints on short-selling and asset price adjustment to private information, *Journal of Financial Economics* 18, 277–311.

Dufour, Alfonso, and Robert F. Engle, 2000, Time and the Price Impact of a Trade, *The Journal of Finance* LV, 2467–2498.

Easley, David, and Maureen O'Hara, 1992, Time and the Process of Security Price Adjustment, *The Journal of Finance* XLVII, 577–605.

Engle, Robert F., 2000, The Econometrics of Ultra-High-Frequency Data, *Econometrica* 68, 1–22.

Engle, Rober F., and Jeffrey R. Russell, 1998, Autoregressive Conditional Duration: A New Model for Irregularly Spaced Transaction Data, *Econometrica* 66, 1127–1162.

Jacobsen, Martin, 2006, *Point Process Theory and Applications*. (Birkhäuser).

Jacod, Jean, and Albert N. Shiryaev, 2002, *Limit Theorems for Stochastic Processes* vol. 228 of *A Series of Comprehensive Studies in Mathematics*. (Springer).

Mainardi, Francesco, Marco Raberto, Rudolg Gorenflo, and Enrico Scalas, 2000, Fractional Calculus and Continuous-time Finance II: the waiting-time distribution, *Physica A* 287, 469–481.

Podlubny, Igor, 1999, *Fractional Differential Equations* vol. 198 of *Mathematics in Science and Engineering*. (Academic Press San Diego, California) 1st edn.

Appendix A. Proofs of propositions and the ML function

Proof Theorem 1.

We show that $(X(t),H(t))$ is described by a stochastic differential equation (SDE), whose coefficients only depend on the process itself. Then it is well known that $(X(t),H(t))$ is a time homogenous Markov process.

In between trades, the backward recurrence time $H(t)$ defined in (6) evolves linearly in t and reverts to zero each time there is a jump in $X(t)$. Therefore $H(t)$ follows the dynamics given by the SDE

$$dH(t) = dt - H(t^-)\,dN_t = dt - \int_{\mathbb{R}_0} H(t^-) z N^1(dt,dz)\,.$$

where $N^1(\omega,dt,dz) = N^1(dt,dz)$ denotes the integer valued random measure that represents the jump measure of the counting process N_t. The intensity of the counting process N_t is given by $u(H(t))$ (see e.g. Jacobsen (2006)) where the hazard function $u(t)$ is given by (2). We can write the predictable compensating measure of $N^1(dt,dz)$ as

$$\nu^1(\omega,dt,dz) = u(H(t))\,dt\delta_1(dz), \tag{A1}$$

where $\delta_1(dz)$ is the Dirac measure centered at 1.

Then it follows that the multivariate dynamics of the two-dimensional process $(X(t),H(t))$ is described by

$$\begin{aligned} \begin{pmatrix} dX(t) \\ dH(t) \end{pmatrix} &= \begin{pmatrix} r-D \\ 1 \end{pmatrix} dt + \begin{pmatrix} 1 & 0 \\ 0 & -H(t^-) \end{pmatrix} \begin{pmatrix} dX(t) \\ dN(t) \end{pmatrix} \\ &= \begin{pmatrix} r-D \\ 1 \end{pmatrix} dt + \int_{\mathbb{R}_0^2} \begin{pmatrix} z_1 \\ -H(t^-)z_2 \end{pmatrix} N^2(dt,dz_1,dz_2) \end{aligned} \tag{A2}$$

where $N^2(\omega,dt,dz_1,dz_2) = N^2(dt,dz_1,dz_2)$ denotes the jump measure of the two-dimensional process $(X(t),N(t))$ on $\mathbb{R}_+ \times \mathbb{R}^2 \setminus \{0\}$. Since the two processes $X(t)$ and $N(t)$ jump at exactly the same times, but with independently distributed jump sizes, the predictable compensator of $N^2(dt,dz_1,dz_2)$ is given by

$$\nu^2(\omega,dt,dz_1,dz_2) = u(H(t))g(z_1)\,dtdz_1\delta_1(dz_2). \tag{A3}$$

Thus the two-dimensional process $(X(t),H(t))$ is described by SDE (A2) with Lipschitz continuous coefficients and predictable compensator that only depend on the process $(X(t),H(t))$ itself (more precisely, on the second component $H(t)$). Then it is well known that $(X(t),H(t))$ is a time-homogenous Markov process.

■

Proof Proposition 1.

We will denote the Fourier transform of a function $g(x)$ by

$$\mathcal{F}\,[g(x)] = \hat{g}(\xi) = \int_{-\infty}^{\infty} e^{ix\xi} g(x)dx\,,$$

where $\xi \in \mathbb{C}$. Hence, assuming the pay-off $G(\cdot)$ is such that we can invert its Fourier transform,

$$\begin{aligned}
V(t) &= e^{-r(T-t)}\mathbb{E}^Q[G(X(T))\,|\,\mathcal{F}_t] \\
&= e^{-r(T-t)}\mathbb{E}^Q\left[\frac{1}{2\pi}\int_{-\infty+i\xi_i}^{\infty+i\xi_i} e^{-i\xi X_T}\hat{G}(\xi)d\xi\,|\,\mathcal{F}_t\right] \\
&= \frac{e^{-r(T-t)}}{2\pi}\int_{-\infty+i\xi_i}^{\infty+i\xi_i} e^{-i\xi\ln S(t)}e^{-i\xi(r-D)(T-t)}\mathbb{E}^Q\left[e^{i\xi\sum_{i=N_t+1}^{N_T}Y_i}\,|\,\mathcal{F}_t\right]\hat{G}(\xi)d\xi\,,
\end{aligned} \tag{A4}$$

where $\mathbb{E}^Q$ denotes the risk-neutral expectation operator.

■

Proof Theorem 3.

We will denote the Laplace transform of a function $f(t)$ by

$$\mathcal{L}\,[f(t)] = \tilde{f}(s) = \int_0^{\infty} e^{st} f(t)dt\,.$$

Further, we assume $H(0) = 0$, i.e. a trade just happened. It will be useful to have an expression for the probability density function $P(n,t)$ of observing n trades during the time interval $[0,t]$. Using the survival function (1) the probability that a trade does not take place before time t is given by

$$P(n=1,t) = \int_0^t \upsilon(s)\Upsilon(t-s)ds = (\upsilon \star \Upsilon)(t),$$

where $\star$ denotes convolution. Then the probability of observing n trades over the interval $[0,t]$ is given by $(\upsilon^n \star \Upsilon)(t)$ and taking its Laplace transform yields

$$\tilde{P}(n,s) = \tilde{\upsilon}(s)^n \tilde{\Upsilon}(s) = \tilde{\upsilon}(s)^n \frac{1-\tilde{\upsilon}(s)}{s}. \tag{A5}$$

Therefore, from Proposition 1, we need to calculate

$$\begin{aligned}
\hat{q}(\xi,0,T) &= \mathbb{E}^Q\left[e^{i\xi\sum_{i=1}^{N_T} Y_i}\right] \\
&= \mathbb{E}^Q\left[e^{(N_T)\Psi(\xi)}\right] \\
\mathcal{L}\left\{\hat{q}(\xi,0,T)\right\} &= \mathcal{L}\left\{\mathbb{E}^Q\left[e^{(N_T)\Psi(\xi)}\right]\right\} \\
&= \mathcal{L}\left\{\sum_0^\infty P(n,T)e^{n\Psi(\xi)}\right\} \\
&= \sum_0^\infty \mathcal{L}\left\{P(n,T)\right\} e^{n\Psi(\xi)} \\
&= \sum_0^\infty \tilde{P}(n,s)e^{n\Psi(\xi)} \\
&= \sum_0^\infty \tilde{\upsilon}(s)^n \frac{1-\tilde{\upsilon}(s)}{s} e^{n\Psi(\xi)} \\
&= \frac{1-\tilde{\upsilon}(s)}{s} \sum_0^\infty \tilde{\upsilon}(s)^n e^{n\Psi(\xi)} \\
&= \frac{1-\tilde{\upsilon}(s)}{s} \frac{1}{1-e^{\Psi(\xi)}\tilde{\upsilon}(s)}.
\end{aligned}$$

where $\tilde{\upsilon}$ is given by (21). Then

$$\begin{aligned}
\hat{q}(-\xi,0,T) &= \mathcal{L}^{-1}\left\{\frac{1-\tilde{\upsilon}(s)}{s}\frac{1}{1-e^{\Psi(-\xi)}\tilde{\upsilon}(s)}\right\} \\
&= E_{\beta,1}\left[-\left(1-e^{\Psi(-\xi)}\right)(T/\tau_o)^\beta\right], \qquad \text{using (A7) below.}
\end{aligned}$$

■

The ML function

In its most general form, the two-parameter Mittag-Leffler function is given by

$$E_{\beta,\gamma}(z) = \sum_{j=0}^{\infty} \frac{z^j}{\Gamma(\beta j + \gamma)}, \qquad \beta > 0, \quad \gamma > 0. \tag{A6}$$

and its Laplace transform, see Podlubny (1999), by

$$\mathcal{L}\left\{t^{\beta n+\gamma-1} E^{(n)}_{\beta,\gamma}(\pm a t^{\beta})\right\} = \frac{n! s^{\beta-\gamma}}{(s^{\beta} \mp a)^{n+1}}, \qquad \mathrm{Re}(s) > |a|^{1/\gamma}, \tag{A7}$$

where $E^{(n)}_{\beta,\gamma}(y) = \frac{d^n}{dy^n} E_{\beta,\gamma}(y)$. This distribution has previously been proposed in the context of financial data in Mainardi, Raberto, Gorenflo, and Scalas (2000).

Appendix B. Empirical and fitted Shifted-Mittag-Leffler survival function

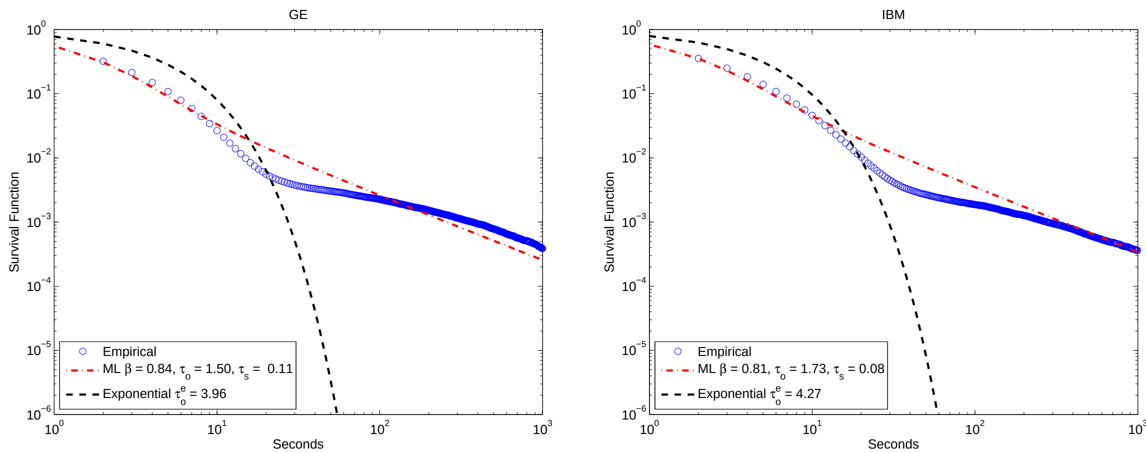

Figure 7. GE and IBM

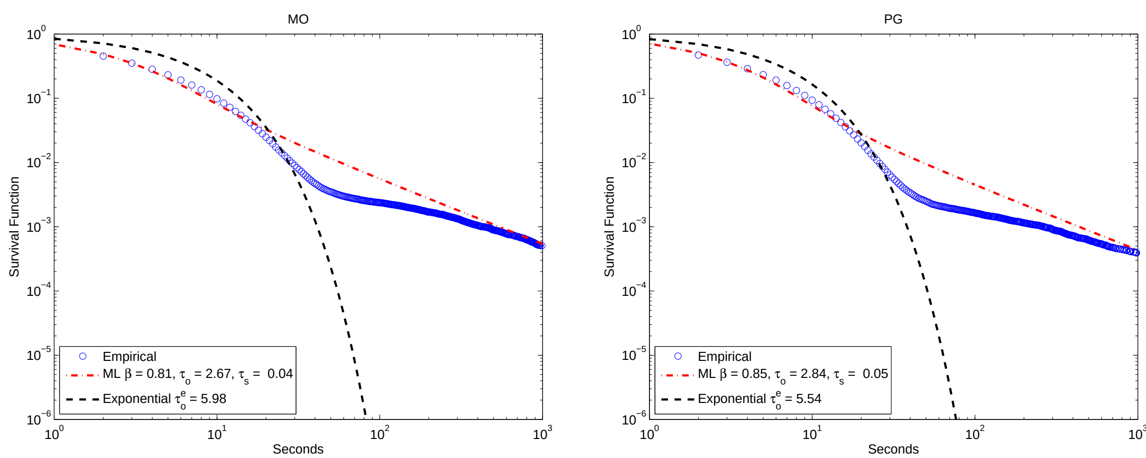

Figure 8. MO and PG

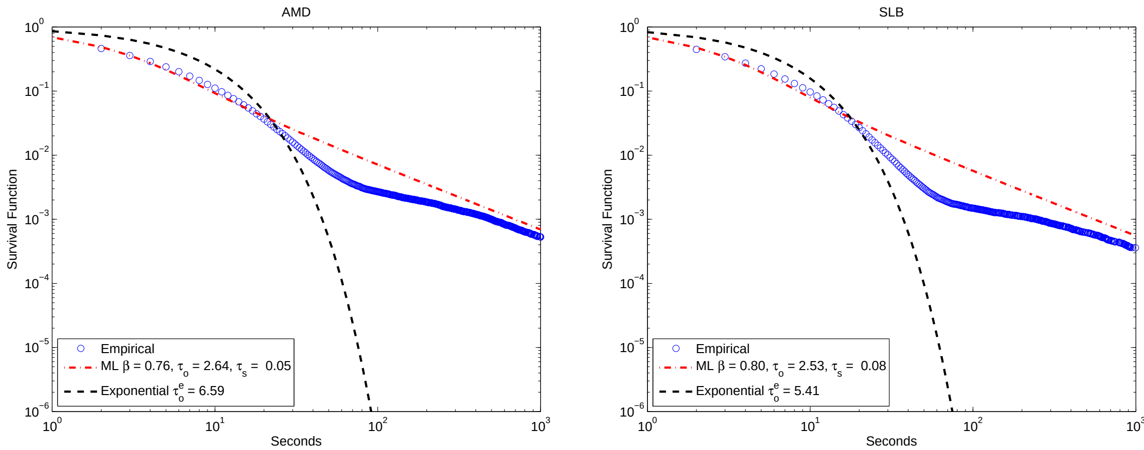

Figure 9. AMD and SLB

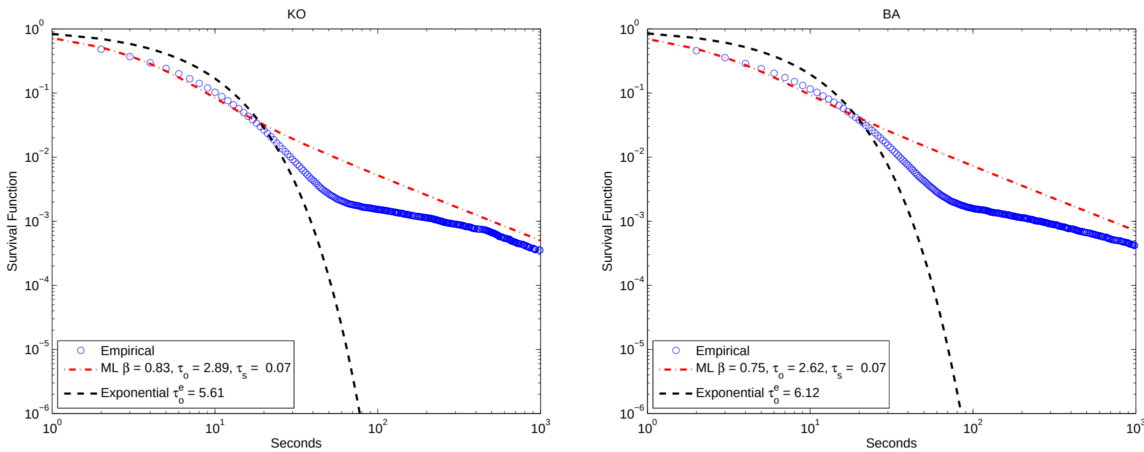

Figure 10. KO and BA

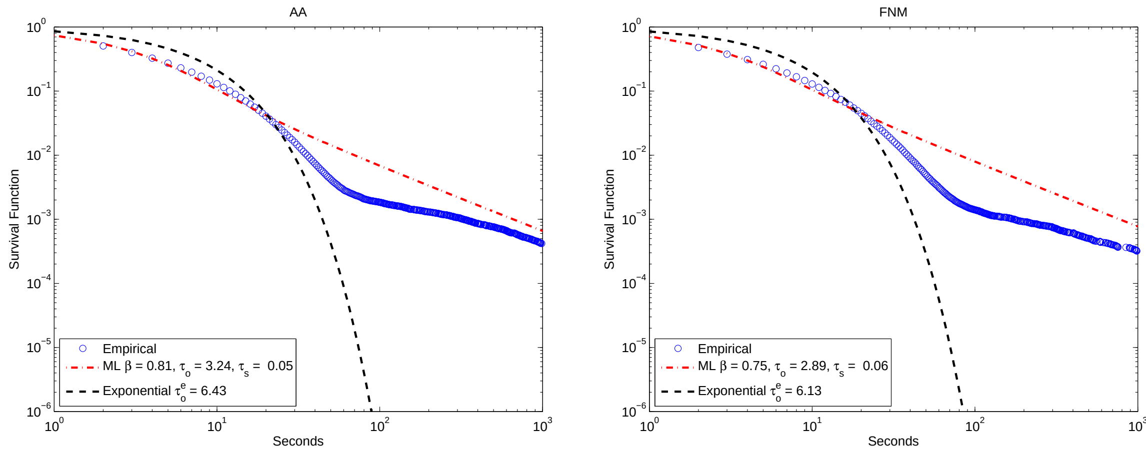

Figure 11. AA and FNM

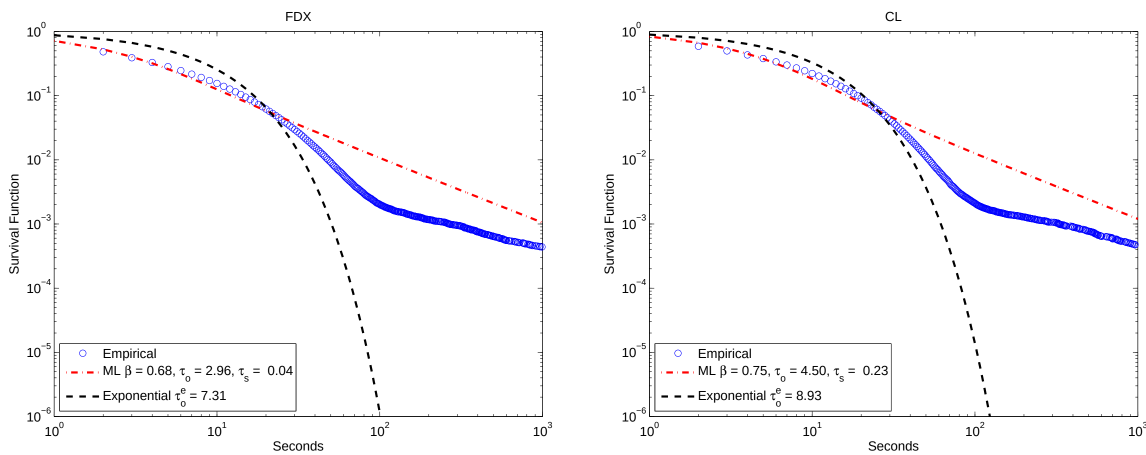

Figure 12. FDX and CL

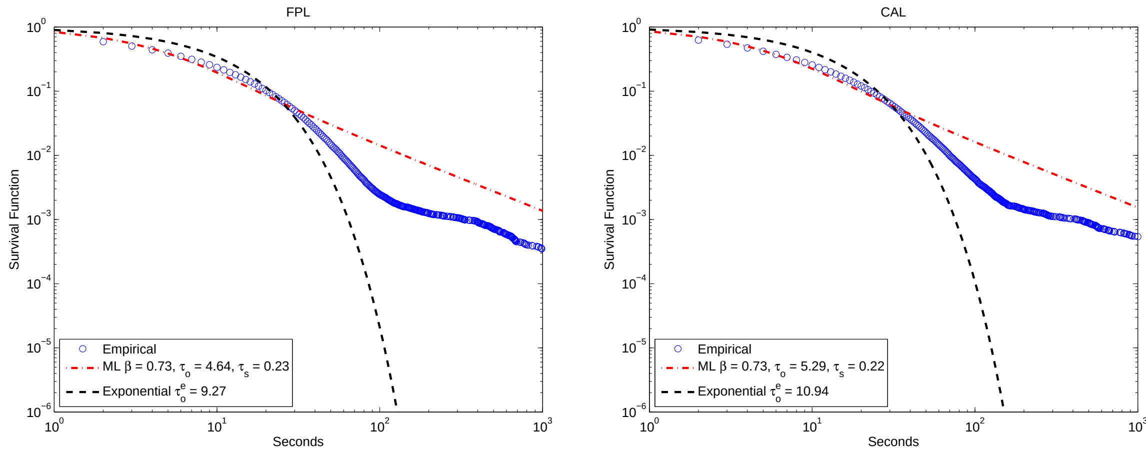

Figure 13. FPL and CAL

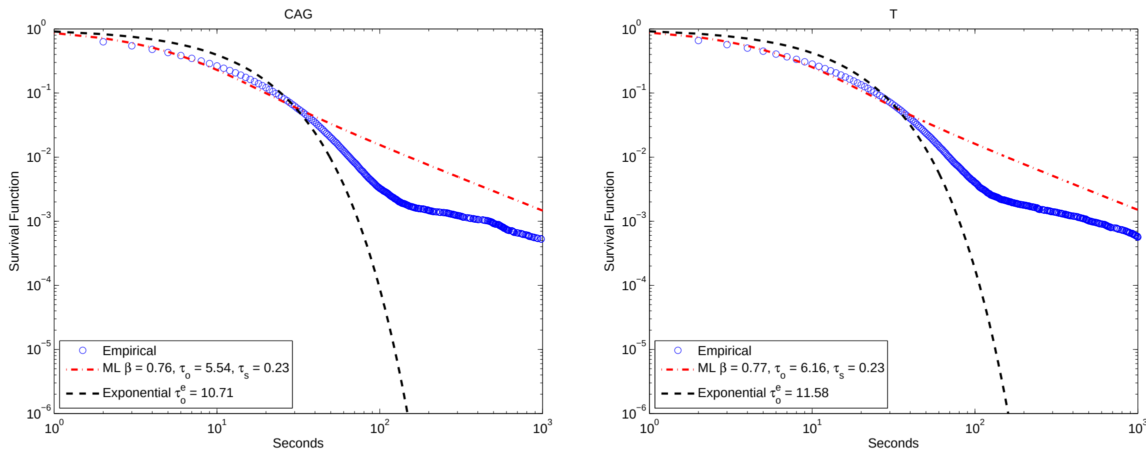

Figure 14. CAG and T

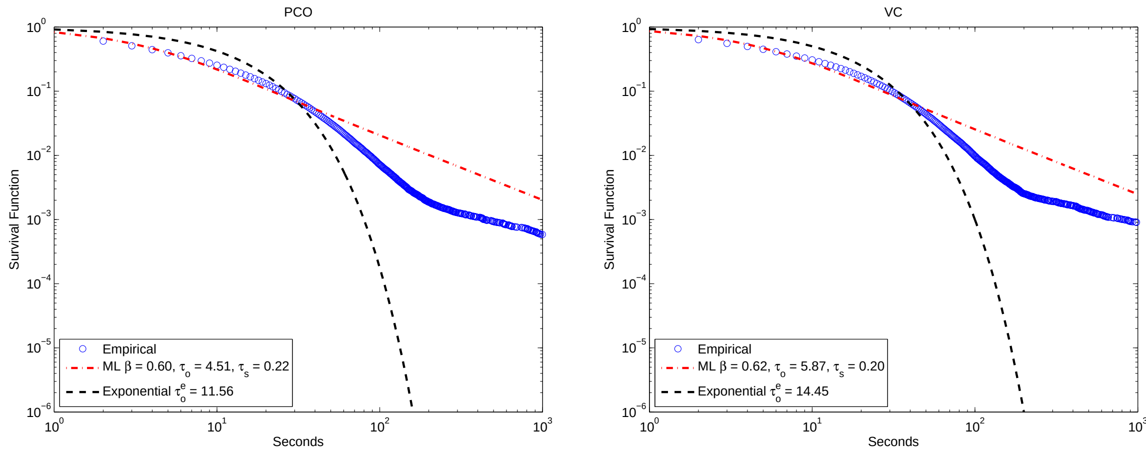

Figure 15. PCO and VC

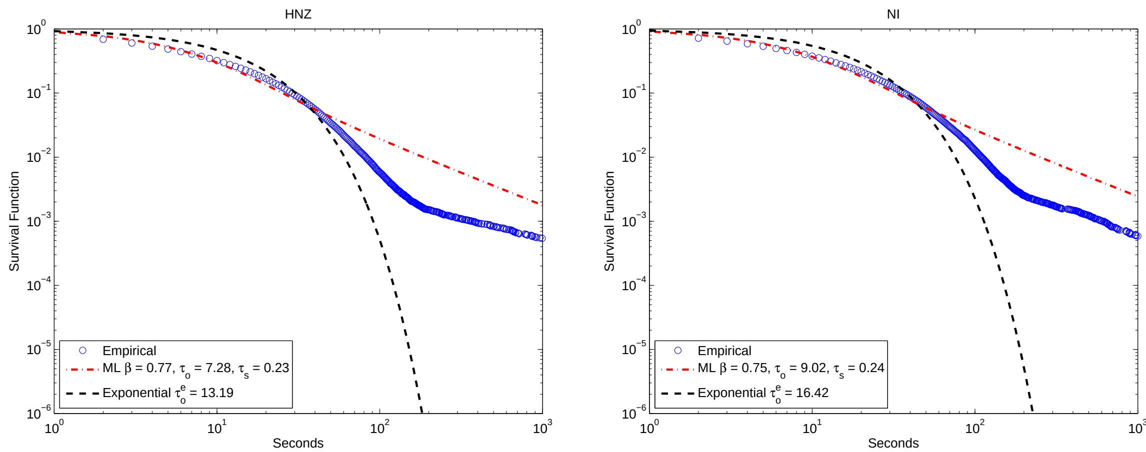

Figure 16. HNZ and NI

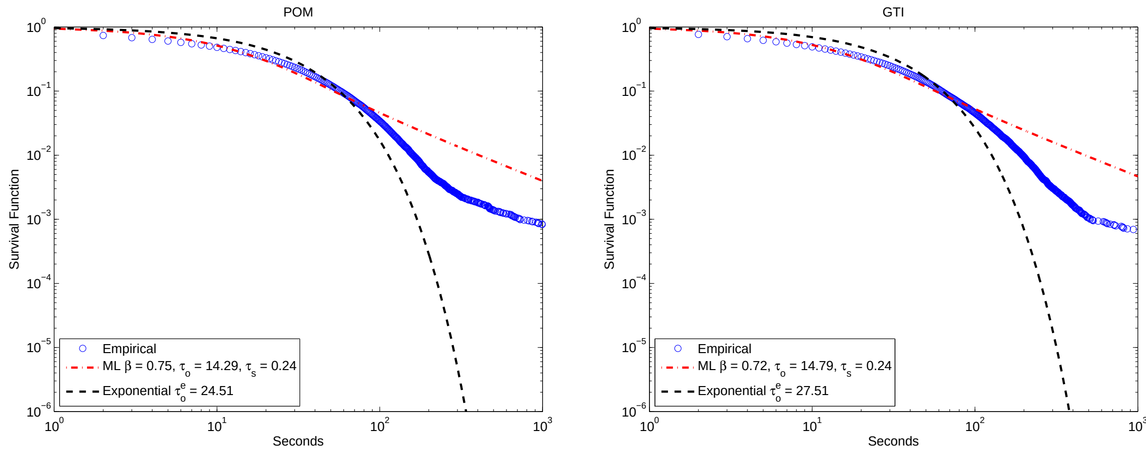

Figure 17. POM and GTI

Appendix C. Calibration of risk-neutral parameters: IBM

Date	Days to Expiry	N	Gaussian Revisions σ	Gaussian Revisions β	FMLS Revisions α	FMLS Revisions σ	FMLS Revisions β
01-Apr	10	6	0.000458	0.717300	1.998	0.000318	0.720044
	10, 35	17	0.000094	0.999999	1.931	0.000051	0.999999
	10, 35, 73	30	0.000108	0.975999	1.916	0.000060	0.964381
	10, 35, 73, 142	45	0.000094	1.000000	1.892	0.000046	0.992243
	10, 35, 73, 142, 185	74	0.000099	0.999996	1.863	0.000039	1.000000
04-Apr	9, 34	16	0.000098	0.999999	1.910	0.000055	0.983908
	9, 34, 72	29	0.000134	0.9473	1.911	0.000082	0.920014
	9, 34, 72, 141	44	0.000111	0.978771	1.888	0.000062	0.948652
	9, 34, 72, 141, 184	73	0.000100	1.000000	1.864	0.000040	1.000000
05-Apr	8, 33	15	0.000099	1.000000	1.889	0.000046	1.000000
	8, 33, 71	29	0.000123	0.961793	1.896	0.000071	0.932939
	8, 33, 71, 140	44	0.000123	0.961808	1.881	0.000060	0.951154
	8, 33, 71, 140, 183	74	0.000101	0.999996	1.861	0.000040	1.000000
06-Apr	7	6	0.000425	0.722593	1.910	0.000254	0.712210
	7, 32	18	0.000100	1.000000	1.884	0.000046	1.000000
	7, 32, 70	32	0.000115	0.972736	1.901	0.000070	0.937586
	7, 32, 70, 139	49	0.000103	0.990980	1.886	0.000060	0.953709
	7, 32, 70, 139, 182	78	0.000101	1.000000	1.867	0.000041	1.000000
07-Apr	6, 31	13	0.000101	1.000000	1.881	0.000046	1.000000
	6, 31, 69	27	0.000104	0.989600	1.903	0.000072	0.937637
	6, 31, 69, 138	42	0.000107	0.985663	1.889	0.000068	0.938441
	6, 31, 69, 138, 181	71	0.000101	1.000000	1.876	0.000044	0.997194
08-Apr	5, 30	14	0.000110	1.000000	1.945	0.000063	1.000000
	5, 30, 68	28	0.000190	0.900856	1.941	0.000116	0.892790
	5, 30, 68, 137	43	0.000173	0.917617	1.915	0.000098	0.904492
	5, 30, 68, 137, 180	72	0.000119	0.979220	1.886	0.000059	0.964613
11-Apr	4, 29	14	0.000540	0.754086	1.955	0.000350	0.746838
	4, 29, 67	29	0.000619	0.729482	1.932	0.000386	0.717918
	4, 29, 67, 136	45	0.000492	0.769292	1.903	0.000284	0.754108
	4, 29, 67, 136, 179	75	0.000322	0.840537	1.874	0.000165	0.823679

Date	Days to Expiry	N	Gaussian Revisions σ	Gaussian Revisions β	FMLS Revisions α	FMLS Revisions σ	FMLS Revisions β
12-Apr	3, 28	13	0.000202	0.919784	1.910	0.000114	0.905947
	3, 28, 66	29	0.000312	0.839863	1.903	0.000180	0.821560
	3, 28, 66, 135	46	0.000308	0.842160	1.884	0.000170	0.819947
	3, 28, 66, 135, 178	76	0.000235	0.887995	1.863	0.000112	0.872700
13-Apr	27	10	0.000759	0.716226	1.927	0.000229	0.816548
	27, 65	25	0.001308	0.627960	1.927	0.000749	0.631227
	27, 65, 134	42	0.000649	0.738271	1.906	0.000363	0.730346
	27, 65, 134, 177	72	0.000315	0.850853	1.887	0.000164	0.837727
14-Apr	26	10	0.001088	0.668316	1.903	0.000197	0.837657
	26, 64	25	0.001877	0.579906	1.925	0.001060	0.586007
	26, 64, 133	42	0.000879	0.699348	1.902	0.000486	0.693781
	26, 64, 133, 176	72	0.000441	0.806794	1.884	0.000230	0.794827
15-Apr	25	9	0.000149	0.998123	1.984	0.000098	0.999732
	25, 63	25	0.000466	0.809579	1.964	0.000293	0.810502
	25, 63, 132	42	0.000474	0.806722	1.943	0.000291	0.801063
	25, 63, 132, 175	72	0.000312	0.872068	1.918	0.000179	0.861563
18-Apr	24	12	0.011902	0.292276	1.796	0.000362	0.698586
	24, 43	28	0.001539	0.620490	1.824	0.000501	0.663351
	24, 43, 62	45	0.000819	0.721936	1.834	0.000327	0.731182
	24, 43, 62, 131	62	0.000480	0.806320	1.833	0.000218	0.789051
	24, 43, 62, 131, 174	92	0.000314	0.872556	1.837	0.000148	0.846786
19-Apr	23	9	0.000487	0.790836	1.932	0.000150	0.890579
	23, 42	22	0.000402	0.823288	1.908	0.000169	0.857384
	23, 42, 61	39	0.000282	0.881639	1.898	0.000127	0.895316
	23, 42, 61, 130	56	0.000267	0.890704	1.889	0.000136	0.880687
	23, 42, 61, 130, 173	86	0.000212	0.926960	1.881	0.000108	0.909810
20-Apr	22	10	0.000951	0.729284	1.878	0.000162	0.895062
	22, 41	24	0.001662	0.637911	1.881	0.000718	0.664969
	22, 41, 60	41	0.001169	0.694911	1.879	0.000559	0.701882
	22, 41, 60, 129	58	0.001001	0.719305	1.870	0.000515	0.709510
	22, 41, 60, 129, 172	88	0.000873	0.740330	1.858	0.000445	0.724647

Date	Days to Expiry	N	Gaussian Revisions		FMLS Revisions		
			σ	β	α	σ	β
21-Apr	21	8	0.002458	0.538957	1.899	0.000311	0.770423
	21, 40	21	0.000682	0.750117	1.897	0.000280	0.785652
	21, 40, 59	38	0.000474	0.809554	1.899	0.000225	0.820470
	21, 40, 59, 128	55	0.000395	0.838556	1.891	0.000208	0.828559
	21, 40, 59, 128, 171	85	0.000305	0.879477	1.878	0.000157	0.862463
22-Apr	20	7	0.001927	0.569939	1.914	0.000334	0.759450
	20, 39	20	0.000431	0.819586	1.889	0.000169	0.854723
	20, 39, 58	37	0.000313	0.873370	1.880	0.000135	0.884846
	20, 39, 58, 127	54	0.000329	0.865097	1.876	0.000165	0.852312
	20, 39, 58, 127, 170	84	0.000249	0.908959	1.866	0.000124	0.888662
25-Apr	19	10	0.027076	0.150392	1.773	0.002712	0.390050
	19, 38	25	0.000713	0.743099	1.821	0.000242	0.769269
	19, 38, 57	42	0.000409	0.835029	1.831	0.000163	0.833948
	19, 38, 57, 126	59	0.000378	0.847751	1.827	0.000172	0.823880
	19, 38, 57, 126, 169	89	0.000311	0.878604	1.827	0.000145	0.848892
26-Apr	18	6	0.000512	0.779657	2.000	0.000560	0.706901
	18, 37	18	0.000310	0.865544	1.949	0.000177	0.873134
	18, 37, 56	34	0.000210	0.931219	1.936	0.000118	0.931968
	18, 37, 56, 125	51	0.000267	0.892632	1.919	0.000153	0.881958
	18, 37, 56, 125, 168	81	0.000242	0.908276	1.897	0.000132	0.891885
27-Apr	17	6	0.000756	0.712010	1.940	0.000233	0.820203
	17, 36	18	0.000316	0.860742	1.938	0.000175	0.867204
	17, 36, 55	34	0.000219	0.922846	1.920	0.011700	0.923192
	17, 36, 55, 124	51	0.000277	0.884841	1.912	0.000156	0.872896
	17, 36, 55, 124, 167	81	0.000245	0.904554	1.897	0.000134	0.888029
28-Apr	16	7	0.001535	0.603984	1.932	0.000434	0.725751
	16, 35	20	0.000407	0.829244	1.914	0.000201	0.842249
	16, 35, 54	37	0.000327	0.866398	1.904	0.000166	0.868008
	16, 35, 54, 123	54	0.000344	0.858393	1.894	0.000185	0.845790
	16, 35, 54, 123, 166	84	0.000309	0.875437	1.885	0.000165	0.857792

Date	Days to Expiry	N	Gaussian Revisions σ	Gaussian Revisions β	FMLS Revisions α	FMLS Revisions σ	FMLS Revisions β
29-Apr	15	6	0.000657	0.735341	2.000	0.000581	0.697424
	15, 34	17	0.000225	0.920759	1.963	0.000139	0.922504
	15, 34, 53	33	0.000188	0.951725	1.942	0.000110	0.949275
	15, 34, 53, 122	50	0.000248	0.906644	1.930	0.000147	0.896451
	15, 34, 53, 122, 165	80	0.000238	0.912866	1.911	0.898226	0.912866
02-May	14	5	0.011543	0.259218	1.919	0.004782	0.325828
	14, 33	18	0.000343	0.853764	1.881	0.000154	0.862809
	14, 33, 52	35	0.000285	0.886290	1.885	0.000139	0.882137
	14, 33, 52, 121	52	0.000268	0.896363	1.878	0.000139	0.878174
	14, 33, 52, 121, 164	82	0.000253	0.905821	1.868	0.000129	0.883145
03-May	13	5	0.001237	0.622567	1.955	0.005310	0.687560
	13, 32	16	0.000294	0.871768	1.940	0.000170	0.871658
	13, 32, 51	32	0.000203	0.935920	1.916	0.000109	0.931191
	13, 32, 51, 120	49	0.000210	0.930860	1.905	0.000116	0.915789
	13, 32, 51, 120, 163	79	0.000197	0.940805	1.887	0.000105	0.920991
04-May	12, 31	15	0.000201	0.931271	1.931	0.000121	0.918467
	12, 31, 50	32	0.000153	0.979814	1.907	0.000084	0.964831
	12, 31, 50, 119	49	0.000186	0.947142	1.895	0.000104	0.925190
	12, 31, 50, 119, 162	79	0.000183	0.950338	1.881	0.000098	0.925661
05-May	11, 30	14	0.000155	0.976139	1.919	0.000101	0.943334
	11, 30, 49	30	0.000137	0.999998	1.901	0.000073	0.985243
	11, 30, 49, 118	47	0.000178	0.957262	1.892	0.000102	0.928288
	11, 30, 49, 118, 161	77	0.000161	0.970512	1.880	0.000100	0.923934
06-May	10	9	0.027844	0.095816	1.600	0.003676	0.253083
	10, 29	22	0.000390	0.825186	1.667	0.000117	0.772995
	10, 29, 48	39	0.000224	0.922345	1.720	0.000083	0.858699
	10, 29, 48, 117	56	0.000236	0.914281	1.743	0.000098	0.848417
	10, 29, 48, 117, 160	86	0.000236	0.914281	1.772	0.000097	0.867299

Table III
IBM risk-neutral parameters April-May 2006

Notes

[1]Below we discuss in detail how consolidated trades from the TAQ database were employed.

[2]In fact, a Kolmogorov-Smirnov test clearly rejects the hypothesis that the data came from an exponential survival function.

[3]In the language of counting processes the process $\sum_{i=1}^{N_t} Y_i$ is called a (0-delayed) renewal process.

[4]Likewise in most stochastic volatility models where the volatility factor under the risk-adjusted measure is essentially the same under the physical measure, but with a linear adjustment.

[5]We calculate this arbitrary duration strictly greater than zero in the following way. Out of all the zero-duration trades we count how many times there were two trades within one second, three trades within one second, etc. Then we calculate a weighted average of number of trades within one second and assume that these occur within 0.5 second instead of deleting them from the sample. Furthermore, for simplicity we do not alter the duration of the trade following those zero-duration trades for which we assigned a non-negative duration.

[6]The constant a could be a function of the parameter β.

[7]To arrive at expression (20) we use the property that

$$\mathcal{L}\{v(t)\} = \mathcal{L}\left\{-\frac{d\Upsilon(t)}{dt}\right\} = -s\mathcal{L}\{at^{-\beta}\} + \Upsilon(0) = -a\gamma(1-\beta)s^{\beta} + 1,$$

and restrict $0 < \beta \leq 1$ to have a valid (monotonic) survival function.

[8]There are a number of articles in the literature that use transform techniques to price and calibrate options, see for example Carr and Wu (2003), Carr and Wu (2004).

[9]Note that we must require $e^{\Psi(-\xi)}$ to be analytic in a line that intersects $[-\infty + i\hat{\xi}, \infty + i\hat{\xi}]$ where $\hat{\xi} > 1$.

[10] We do not calibrate to lots where there are less than 5 options.

[11]For simplicity we assumed that τ_o remained the same under both the physical and statistical measure (see Table II for IBM) and that $\tau_s = 0$.

[12]Figure 3 does not include the case $\beta = 1$ where IV becomes 0.30 for all expiries as expected. However, the case $\beta = 1$ can be seen in Figure 4.